Tec Empresarial
P-ISSN: 1659-2395; E-ISSN: 1659-3359

GRID SEARCH BASED HYPERPARAMETER OPTIMIZATION FOR PREDICTING CORONARY ARTERY DISEASE USING MACHINE LEARNING TECHNIQUES

CMM. Mansoor
Department of Infromation Technology, South Eastern University of Sri Lanka, Sri Lanka
mansoorr@seu.ac.lk

Sarat Kumar Chettri
Department of Computer Applications, Assam Don Bosco University, India
sarat.chettri@dbuniversity.ac.in

HMM.Naleer
Department of Computer Science, South Eastern University of Sri Lanka, Sri Lanka
hmmnaleer@seu.ac.lk

Abstract— Heart disease is a major worldwide health concern that raises rates of morbidity and death. To successfully prevent and manage Coronary Artery Disease (CAD), early and precise prediction is essential. Still, it's a difficult undertaking to do. In order to forecast cardiac illness, this paper suggests a machine learning model that makes use of a number of preprocessing stages, hyperparameter optimization strategies as well. Making more accurate judgments for cardiac problems is facilitated by the prediction of this type of cardiac sickness. Hence, one practical method for predicting the illness as soon as feasible is to employ Grid Search Optimization (GSO) Machine Learning (ML) models. The GSO is used to find the optimal values of the hyper-parameters. The most advanced technique involves choosing the feature and adjusting the hyperparameter in tandem with the model search to reduce the false-negative rate. Feature selection for multivariate analysis based on statistical and correlation matrix analysis. Extensive comparative results are generated for the Grid Search models using retrieval, F1 score, and accuracy metrics. The performance results of the machine learning techniques were obtained. The Random Forest (RF) with tuned hyperparameters achieved the highest accuracy of 90.14% for predicting the disease. Z-Alizadeh sani dataset was used for this research. The dataset was pre-processed to making it suitable for a Machine Learning (ML) model. The Synthetic Minority Oversampling Technique (SMOTE) as well as SelectKBest techniques were used to improve a model's effectiveness and precision. SMOTE often results in a higher recall at the expense of decreased accuracy. The application of grid search to improve performance in training and testing was innovative in this study. The outcomes of the experiments were used to determine the optimal parameters for heart disease prediction. The methodology employed in this work greatly surpassed the findings of previous research that focused on the prediction of heart disease, according to comparative analyses that were also conducted.

Keywords—Machine Learning Algorithm, Hyperparameter, Heart Disease Prediction, Gridsearchcv

Introduction

Maintaining general health and well-being depends on the heart, an essential organ that pumps blood throughout the body. Preserving a healthy cardiovascular system is crucial as compromised heart function can impair other body parts such as the brain and kidneys. Cardiovascular disease is a serious global health problem as it is a common and deadly illness that causes millions of deaths annually. This kind of cardiovascular illness has become more prevalent recently. The World Health Organisation (WHO) [1] estimates that heart disease accounts for around 31% of all deaths worldwide, with over 23.6 million individuals at risk of passing away from the illness. As a result, physicians insist that patients with Coronary Artery Disease (CAD) be identified based on medical tests and diagnosis as well as symptoms [2].

Machine Learning (ML) is being used considerably more in healthcare diagnosis. This might be partly explained by improvements in the identification and categorization of illness, data that helps physicians find and diagnose illnesses, preservation of human health, and a drop in mortality rates. These days, ML facilitates prediction and decision-making. The primary driving force behind the research is the utilization of grid search models to enhance classificator performance. The grid search for the best parameters yields the hyperparameter. Key elements are selected from the grid search procedure, and suitable features are selected for the model's development. With these feature choices, target classes are expected.

It is still difficult to forecast CAD with any degree of reliability, despite prior scientific efforts. The early diagnosis of coronary artery disease using clinical indicators is lacking in studies. Daily blockages of the heart caused by plaque diminish the blood flow to the heart. This phase of CAD is known as the early stage. Plaque has substance-filled walls made of cholesterol. Gradually, less oxygen is needed for cardiac function. The wall and valve will contract during this period, giving the reference image an odd aspect. With a better degree of accuracy, our suggested approach is utilized to anticipate CAD early.

This paper fills this vacuum by presenting a prediction model for CAD that includes preprocessing approaches, hyperparameter optimization strategies. The prediction of this kind of cardiac illness aids in the making of more precise assessments for cardiac issues. Thus, using Grid Search Optimization (GSO) Machine Learning (ML) models is a useful way to forecast the sickness as soon as possible. The hyper-parameters' ideal values are discovered using the GSO. The most sophisticated method lowers the false-negative rate by selecting the feature and modifying the hyperparameter simultaneously with the model search. Using statistical and correlation matrix analysis to guide feature selection for multivariate analysis. For the Grid Search models, extensive comparison data are produced utilizing accuracy, retrieval, and F1 score criteria. By comparing the classification accuracy of many ML methods, such as Random Forest (RF), K-neighbor Classifier (KNN), and Support Vector Machine (SVM) the model seeks to precisely forecast CAD. To improve the performance of the model, additional preprocessing procedures are used, such as addressing missing values, eliminating outliers as well as duplicates, and normalizing data.

The rest of this article is articulated as follows: The relevant research on earlier studies used in the ML methods used in the heart disease prediction system is covered in Section 2. More information about the planned work's methodology is provided in Section 3. The experimental results, comparative analysis of the earlier research, and methodology are covered in Section 4. Our results and recommendations for further study are presented in Section 5.

Literature Review

The heart, the most significant part in the human body, flows blood through the whole body. The arteries, veins, and capillaries that make up the circulatory system are related to the heart. The heart is disposed to disease and injury even though it is one of the most energetic organs in the human body [3]. Because of the danger it brings to human life and the illnesses and injuries it causes, its existence cannot be ignored. Cardiac illness affects the heart's pumping processes, rendering them inefficient. Heart problems are the cause of a class of diseases known as CAD. CAD symptoms include fatigue, swollen feet, muscular weakness, as well as shortness of breath. Thus, it is essential that the heart functions normally [4]. Heart conditions are increasingly ranking among the world's foremost causes of mortality. Consequently, a large number of scholars worldwide started concentrating on using the extensive databases to predict cardiovascular conditions [5].

There is a connection between electrocardiography (ECG), which measures the heart's electrical activity, and CAD. Because of its low cost, wide operational availability, and non-invasive nature, ECG should be looked at as a potential screening tool [6]. However, significant inter- and intra-individual variability may make it challenging to reliably screen for CAD patients using patient ECG recordings [7]–[9]

An ensemble model for the diagnosis of CAD [10] was built using three neural network models [11]. The performance was assessed using SAS Enterprise Miner 5.2. The performance did not improve even when the number of NN was doubled. A cascaded neural network, a self-organizing network, and a SVM with RBF function were used to train and evaluate 270 patient records. A ML model called Naïve Bayes was created to forecast CAD [12].

By using them to the given dataset, scholars may evaluate how well contemporary architectures like LSTM and GRU perform when analyzing patients' cardiac issues. The most precise technique for identifying CAD is, in certain respects, coronary angiography [13]. The images from coronary angiography can be used to determine the degree of arterial stenosis and evaluate the severity of CAD and ACS and other conditions. But it's an expensive and time-consuming process [14].

The personalized data of the patient is substituted with some fictitious values from a security perspective. Unlike previous ML approaches, the study takes into account and uses medical database performances for predicting heart disease [15]. A basic but significant probabilistic model is the Naive Bayes. In this note, it will serve as an ongoing example. Specifically, in the case of completely seen data, we will first investigate maximum-likelihood estimation; in the case of partially observed data, when example labels are lacking, we will next consider the expectation maximization (EM) approach. The Naive Bayesian classifier relies on the independence assumptions between predictors and the Bayes theorem. Large datasets benefit greatly from the simplicity and lack of complex iterative parameter estimation found in naive Bayesian models. The Naive Bayesian classifier is popular because it frequently outperforms more complex classification techniques, even if it is incredibly simple. Regarding the Naive Bayesian model, the method promises to be highly significant and successful in handling categorization, akin to ML [16], [17]. However, the naive Bayes model presumes that every prediction, or characteristic, is independent, which is rarely the case in reality. This restricts the algorithm's usage in practical scenarios [18].

The method of hyperparameter tweaking to find the best values for a particular model is called GridSearchCV. As previously stated, hyperparameter values have a major impact on a model's performance. To achieve high precision and find the ideal value for the hyperparameters, hyperparameter tuning is carried out.. The fit method of the Scikit-learn GridSearchCV class offers a grid of tweaking classification methods. It makes it possible to train any ML algorithm in a single, reliable environment and to modify the corresponding hyperparameters. After the hyperparameters have been adjusted to an appropriate level, the full training dataset is utilized to create an accurate model [19].

Methodology

This section begins with a summary of the resources (such as the dataset) that were used in our study. After that, a brief explanation of the feature engineering methods we took into account for our research is provided.

System architecture diagram

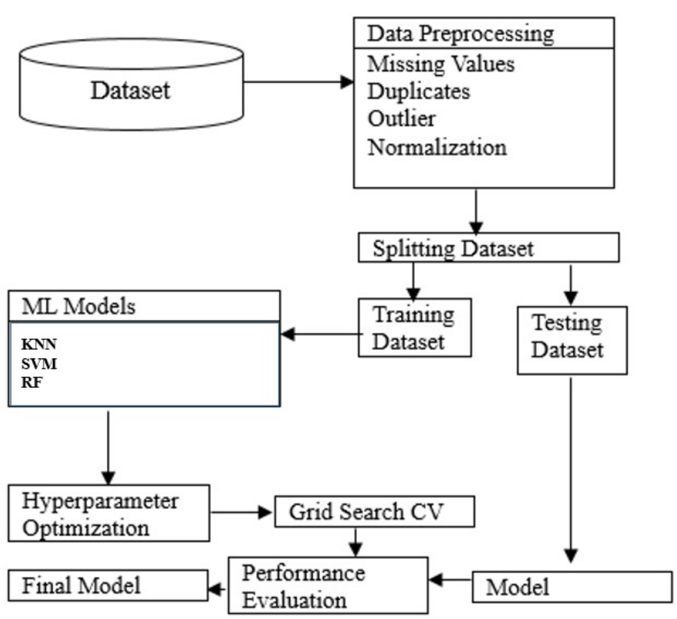


Fig 1: The proposed model developed to predict the heart disease

An architecture diagram outlining the sequential stages needed in estimating the likelihood of CAD is shown in Fig. 1. The data pretreatment steps such as cleaning, eliminating duplicate entries, finding outliers, and scaling data are graphically shown in the diagram. After a comprehensive examination of the dataset to ensure that no missing data existed, it was determined that there were no missing values. In order to train and optimize the model, it also highlights the use of the hyperparameter tuning procedure. The flow chart facilitates understanding of the research process by providing a visual depiction of the study's methods.

Data Collection

To improve our analysis in this study, we integrated three Z-Alizadeh Sani dataset [20]. The 303 almost completed items in the Sani dataset were used to build the proposed approach. The dataset's substantial utilization rate makes it useful for comparing the precision of detection findings with those from other research projects. Twelve characteristics with 303 instances have been chosen.

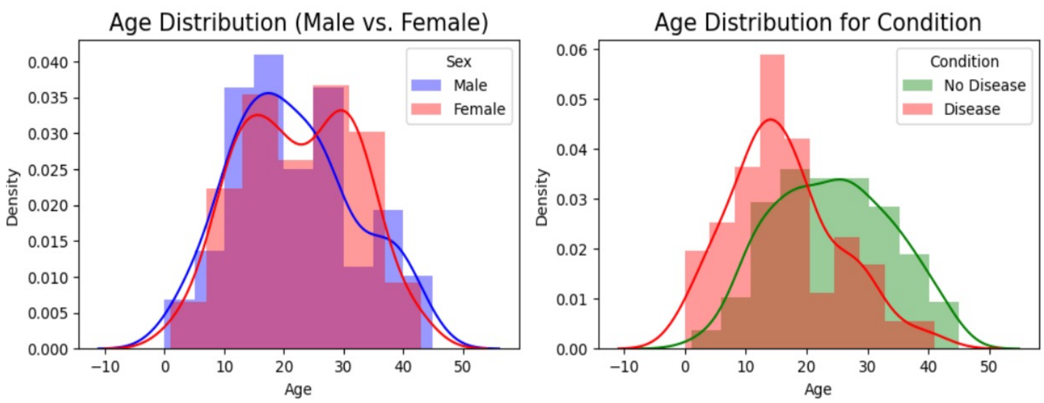
CAD dataset [21]

Attribute	**Description**	**Type**	**Range**	**Mean**
Age	Years	Numeric	30-86	58.90
DM	Diabetes Mellitus	Numeric	Yes,No	0.30
HTN	Hypertension	Numeric	Yes,No	0.60
Current Smoker	-	Numeric	Yes,No	0.20
Typical Chest Pain	-	Int64	Yes,No	0.54
Atypical	-	Object	Yes,No	-
Q Wave	-	Int64	Yes,No	0.05
St Elevation	-	Int64	Yes,No	0.05
St Depression	-	Int64	Yes,No	0.23
Tinversion	-	Int64	Yes,No	0.30
Poor R Progression		Object	Yes,No	-
Region RWMA		Int64	0-4	0.62
Target Class:Cath		Object	CAD,Normal	-

Preprocessing

Initially the dataset was thoroughly examined to make sure there were no instances of missing data, and the results showed that there were no missing values. Secondly, we changed the format by using Label Encoder. The Label Encoder transforms the intended labels into numerical representations so that machines can understand them and diagnose illnesses more accurately. Next, we employed the approach of SMOTE for data balancing. The SMOTE approach was used to balance the data sets. Furthermore, a careful review of duplicate values was done to guarantee data consistency because the dataset was created by combining three different datasets. The lack of duplicates was verified by this study, supporting the dataset's accuracy. To find any extreme numbers that may potentially distort the data, an outlier analysis was also carried out. Interestingly, there were no outliers found, demonstrating the dataset's resilience. The data's analysis and interpretability were improved by this scaling procedure. The comprehensive data preparation procedures laid a strong basis for the trustworthy and excellent analyses in this research.

Age Distribution



Age distribution with and without CAD condition in the dataset

The age distribution with and without the CAD condition is displayed in the fig 2. The age categorized between 0 to 50. Comparatively large amount of male patients were founded with CAD condition than the female. Also the age between 10 to 40 were identified as highest CAD condition on patients.

Modeling

KNN

Often referred to as KNN or k-NN, the k-nearest neighbours method is a non-parametric supervised learning classifier that groups individual data points together for categorization or prediction based on closeness. It may be applied to regression or classification issues as well, although it is usually employed as a classification strategy based on the idea that equivalent points may be put near to each other.

This distance metric, which is exclusively applicable to real-valued vectors, is the most often used one. It measures a straight line between the query location as well as the other point that is measured.

SVM

The support vector machine (SVM) is a ML algorithm that determines boundaries between data points based on predefined classes, labels, or outputs. It uses supervised learning models to solve complex problems related to classification, regression, and outlier detection. SVMs are widely used in many sectors, including speech and image recognition, natural language processing, healthcare, and signal processing applications.

RF

RF refers to Random Forest classifier. The popular ML algorithm RF, which aggregates the output of many decision trees to produce a single outcome. Its versatility and ease of use, combined with its ability to handle both regression and classification issues, have driven its popularity.

Correlation Matrix

Correlation score with target variable

The concept of correlation describes the connections between one or more variables. These factors might be characteristics of the raw data that were utilized to forecast our target variable. A statistical approach called correlation shows how one variable changes or moves in connection to another one. It provides us with a general understanding of how closely the two variables are related. This bi-variate analysis measure explains the relationship between many variables. The correlation matrix, which represents the correlation coefficients between sets of variables, is seen in Fig.3.

A heat map is a visual aid that illustrates the proportion of each feature's association with each other as well as the correlation between features. Additionally, it explains how each attribute correlates with the desired feature. A collection of computing techniques called statistics are used to evaluate unprocessed data and turn it into understandable information. It's one of the instruments in the ML toolbox. ML and statistics are two disciplines that are closely connected. In order to determine the characteristics of common and associated data samples, such as mean, standard deviation, max and min values, and descriptive statistics were computed on the dataset. The goal variable, heart disease, has the largest positive connection with St Elevation out of all the other variables, as can be seen from the correlation plot and heatmap above. With a correlation value of 0.14 and -0.72, respectively, the variables are still categorized as having moderate overall association. As a result, using this dataset for analysis is deemed safe. There is no need to remove any columns at this time.

Hyperparameter Tuning Optimization

Hyperparameter optimization, or tuning in ML, is the task of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process. Numerous optimization techniques exist, each having advantages and disadvantages. It is common to discover the values of other parameters. Selecting the optimal hyperparameters has a noteworthy impact on the performance model. The optimal hyperparameter combination was found through experiments using a variety of optimization approaches, and it was then applied in three ML algorithms: the RF classifier, the support vector machine classifier, and the K-nearest neighbour classifier. One of the optimization issues is the careful tweaking of ML algorithms. In hyperparameter optimization, the GridSearchCV method is widely applied to overcome obstacles and boost model accuracy. A tried-and-true technique that takes into account all hyperparameter combinations is GridSearchCV. GridSearchCV uses the learning rate and layer count as hyperparameters.

Evaluation Metrics

The performance criteria included in this study were essential in assessing the classifiers' efficacy and accuracy. AUC-ROC, F1-score, accuracy, recall, precision, and recall were among the measures used. These metrics provide insightful information about several facets of the classifiers' operation. Table 2's confusion matrix, which presents the classification findings in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), served as the basis for the evaluation. properly anticipated cases of the positive class are represented by TP, and properly predicted instances of the negative class are represented by TN. examples that were mistakenly forecasted as the negative class are designated as FN, whereas examples that were mistakenly projected as the positive class are designated as FP.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{1}$$

$$Precision = \frac{TN}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FP} \tag{3}$$

$$F - score = 2\,x \frac{TP}{TP+FP} \tag{4}$$

Confusion matrix

Actual Class	Predictive Class	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

AUC-ROC

A measure of a classification model's effectiveness over all potential classification thresholds is the area under the ROC curve, or AUC-ROC. We may transfer the sigmoid output of a binary classification to a binary category in machine learning by using the classification threshold, also known as the decision threshold. As the categorization threshold changes from 0 to 1, it illustrates the trade-off between the genuine positive rate and the false positive rate. A perfect model has an AUC-ROC value of 1, whereas a mediocre model has an AUC-ROC value of 0.5. PRC: In a binary classification issue, a precision–recall curve (PRC) is a graphical representation that shows the correlations between precision and recall for various classification levels. It demonstrates how adjustments to the categorization threshold affect the accuracy vs. recall trade-off.

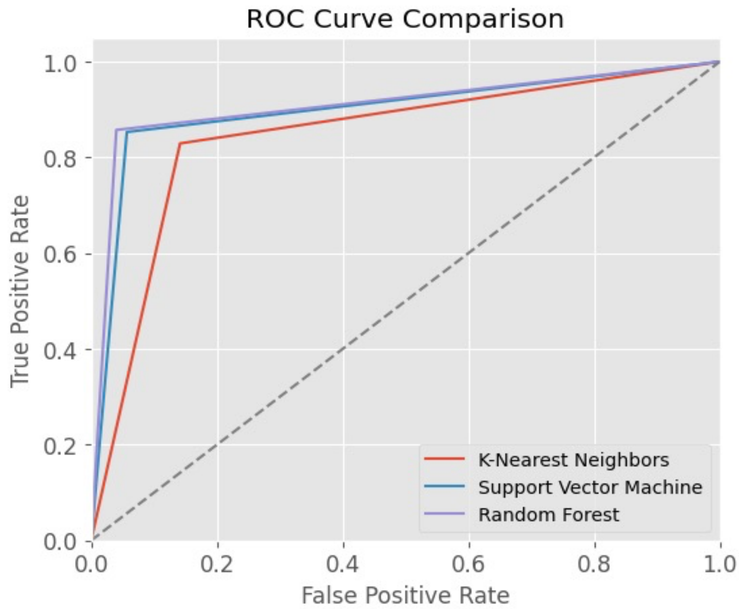
Results and Discussion

We used preprocessing processes, hyperparameter optimization strategies, and a variety of classification algorithms in our experimental investigation on a CAD dataset. The outcomes showed that the algorithms could predict CAD illness with a reasonable degree of accuracy, with some of them outperforming others. The enhanced performance was a result of the preparation procedures, which included resolving missing values, eliminating duplicates and outliers, and normalizing the data.

The most appropriate method for precise classification was determined by comparing the three algorithms that were used: KNN, RF, and SVM. Python programming was used to do the analysis after the dataset was divided into training and testing segments using an 80:20 ratio.
Performance Evaluation Using Default Hyperparameter Settings
performance metrics for the algorithms using the default hyperparameter settings.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
KNN	86.08	85.17	87.14	85.79
SVM	88.35	88.02	89.47	88.18
RF	91.24	90.79	92.13	91.02

With the default hyperparameter settings, the TABLE III shows the overall accuracy, precision, recall, and F1-score values for all three algorithms. At 91.24%, the RF classifier had the best accuracy. Additionally, it generated the greatest F1-score of 91.02, recall of 92.13%, and accuracy of 90.79%. When considering the KNN, it gained 86.08% of accuracy, 85.17 of precision, 87.14 of recall and 85.79% of F1-score. Then, SVM has got 88.35% of accuracy, 8802% of precision, 89.47% of recall and 88.18% of F1-score. We used a ROC curve, as shown in Fig. 4, to further confirm our findings.



ROC curve comparision with default hyperparameter

Performance Evaluation Using Grid Search CV

Based on their higher performance as measured by 5-fold CV, these hyperparameters were chosen. The effect of these hyperparameters on the model's performance is seen in Table IV. The RF classifier achieved the highest.

The optimal hyperparameters for the algorithms as the result of hyperparameter optimization using grid search CV

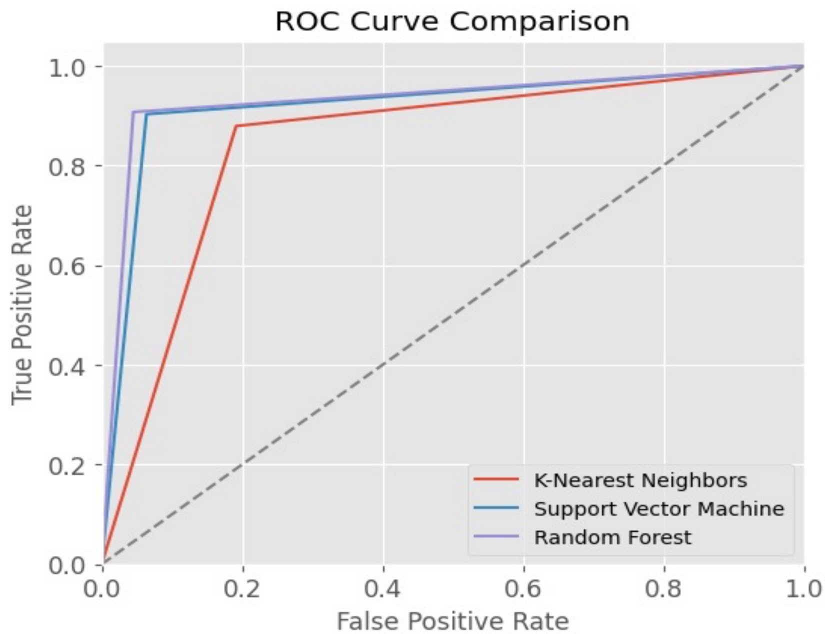
Algorithm	**Optimal Hyperparameters**
KNN	leaf_size: 5; metric: 'minkowski'; n_neighbors': 8; p: 1, weights': 'uniform'
SVM	C=20 , degree = 3 ; gamma = 'auto' ; kernal='rbf',class_weight=None;cache_size=200;random_state=None
RF	Bootstrap=True;criterion='gini;max_depth=10;max_features='sqrt';min_samples_split=2; min_samples_leaf=4; n_estimators=100,class_weight='None;

The performance metrics for the algorithms using the optimal hyperparameters according to the grid search CV

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
KNN	86.93	85.74	87.56	85.95
SVM	88.86	88.14	90.83	88.43

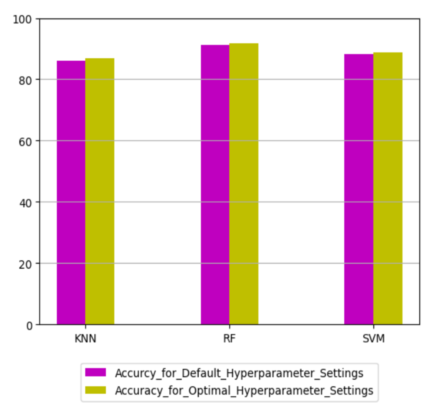
RF	91.87	91.14	92.75	91.58

The table V displays the overall accuracy, precision, recall, and F1-score values for each of the three methods using the optimal hyperparameter settings. The RF classifier achieved the highest accuracy, at 91.87%. It also produced the highest accuracy of 91.14%, recall of 92.75%, and F1-score of 91.58%. KNN has got 86.93% in accuracy, 85.74% in precision, 87.57% in recall, and 85.95% in F1-score. Afterwards, SVM's results include 88.86% accuracy, 88.14% precision, 90.83% recall, and 88.43% F1-score. As seen in Fig. 5, we employed a ROC curve to further corroborate our results.



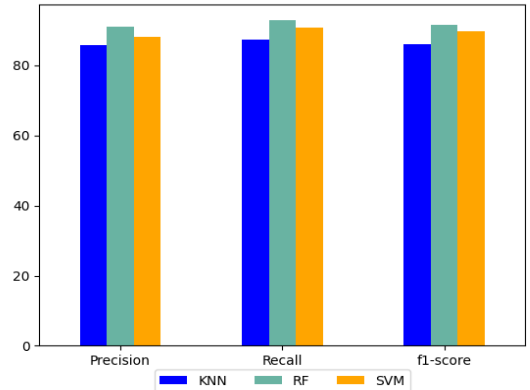
ROC curve comparision with optimal hyperparameter Grid search

Performance Comparison



Performance comparision of each classifier

When comparing the two performance values RF hits the highest score. Then, SVM has got the average of 88%. Finally KNN got lowest score among other classifiers.



Comparision of diffreent classifers

In overall these research articles increased their precision, recall and F1-score level by using three classifiers or algorithms. In order to that RF is achieved highest score of precision, recall and F1-score. Secondly, SVM got the high scores. Finally, KNN has got the scores in third level. The statistic used to evaluate how well models perform in classification tasks. In this regard, our innovative method RF classifier achieved with the best accuracy, recall, precision and F1-score.

Conclusion

The heart, a vital body part accountable for pumping blood throughout the body, is critical to keeping overall health and wellbeing. It is important to keep the cardiovascular system in good working order since problems with the heart can affect other organs like the brain and kidneys. Since cardiovascular disease is a widespread and fatal ailment that claims millions of lives each year, it poses a severe threat to world health. These days, decision-making and prediction are made easier by ML. The use of grid search models to improve classificator performance is the main motivation for the research. The hyperparameter is found by grid search for optimal parameters. From the grid search process, important components are chosen, and characteristics that are appropriate for the model's development are chosen. Target classes are expected with these feature selections.

Our studies' findings showed that the algorithms' performance was greatly impacted by the hyperparameter optimization strategies. In particular, our suggested model's improved accuracy of 91.87% was attained by applying the random method, an 80:20 split ratio, and the ideal hyperparameters found by grid search CV. Researchers and practitioners working on heart disease-related activities may find this model useful.

References

[1] H. Ayatollahi, L. Gholamhosseini, and M. Salehi, “Predicting coronary artery disease: A comparison between two data mining algorithms,” BMC Public Health, vol. 19, no. 1, pp. 1–9, 2019, doi: 10.1186/s12889-019-6721-5.

[2] C. N. and S. R. S. Prabu, B. Thiyaneswaran, M. Sujatha, “Grid Search for Predicting Coronary Heart Disease by Tuning Hyper-Parameters,” Computer Systems Science & Engineering, 2022.

[3] T. Suresh, T. A. Assegie, S. Rajkumar, and N. K. Kumar, “A hybrid approach to medical decision-making: diagnosis of heart disease with machine-learning model,” International Journal

of Electrical and Computer Engineering, vol. 12, no. 2, pp. 1831–1838, 2022, doi: 10.11591/ijece.v12i2.pp1831-1838.
[4] O. Y. Atkov et al., "Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters," Journal of Cardiology, vol. 59, no. 2, pp. 190–194, 2012, doi: 10.1016/j.jjcc.2011.11.005.
[5] A. D. Jamthikar, D. Gupta, L. E. Mantella, L. Saba, A. M. Johri, and J. S. Suri, "Ensemble Machine Learning and Its Validation for Prediction of Coronary Artery Disease and Acute Coronary Syndrome Using Focused Carotid Ultrasound," IEEE Transactions on Instrumentation and Measurement, vol. 71, 2022, doi: 10.1109/TIM.2021.3139693.
[6] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network," Knowledge-Based Systems, vol. 132, pp. 62–71, 2017, doi: 10.1016/j.knosys.2017.06.003.
[7] J. Chen, A. Valehi, and A. Razi, "Smart Heart Monitoring: Early Prediction of Heart Problems through Predictive Analysis of ECG Signals," IEEE Access, vol. 7, pp. 120831–120839, 2019, doi: 10.1109/ACCESS.2019.2937875.
[8] G. Phadke, M. R. Rajati, and L. Phadke, "Prediction of coronary artery disease using electrocardiography: A machine learning approach," Proceedings - International Conference on Machine Learning and Cybernetics, vol. 2020-Decem, pp. 175–180, 2020, doi: 10.1109/ICMLC51923.2020.9469585.
[9] F. Sanfilippo and C. Pacchierotti, "A wearable haptic system for the health monitoring of elderly people in smart cities," International Journal of Online Engineering, vol. 14, no. 8, pp. 52–66, 2018, doi: 10.3991/ijoe.v14i08.8571.
[10] E. Araka, R. Oboko, E. Maina, and R. Gitonga, "Using Educational Data Mining Techniques to Identify Profiles in Self-Regulated Learning: An Empirical Evaluation," International Review of Research in Open and Distance Learning, vol. 23, no. 1, pp. 131–162, 2022, doi: 10.19173/IRRODL.V22I4.5401.
[11] N. Sinha, T. Jangid, A. M. Joshi, and S. P. Mohanty, "iCardo: A Machine Learning Based Smart Healthcare Framework for Cardiovascular Disease Prediction," 2022, [Online]. Available: http://arxiv.org/abs/2212.08022.
[12] E. Nasarian et al., "Association between work-related features and coronary artery disease: A heterogeneous hybrid feature selection integrated with balancing approach," Pattern Recognition Letters, vol. 133, pp. 33–40, 2020, doi: 10.1016/j.patrec.2020.02.010.
[13] L. Verma, S. Srivastava, and P. C. Negi, "A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data," Journal of Medical Systems, vol. 40, no. 7, 2016, doi: 10.1007/s10916-016-0536-z.
[14] A. H. Shahid, M. P. Singh, B. Roy, and A. Aadarsh, "Coronary artery disease diagnosis using feature selection based hybrid extreme learning machine," Proceedings - 3rd International Conference on Information and Computer Technologies, ICICT 2020, pp. 341–346, 2020, doi: 10.1109/ICICT50521.2020.00060.

[15] Z. F. Khan and S. R. Alotaibi, “Applications of Artificial Intelligence and Big Data Analytics in m-Health: A Healthcare System Perspective,” Journal of Healthcare Engineering, vol. 2020, 2020, doi: 10.1155/2020/8894694.

[16] S. A. Asra, “Various Technologies used in Smart Transportation System in the Direction of a Smart City : A Systematic Review,” 2022.

[17] S. A. Asra, “Applications of mobile ad hoc network (MANET): a survey,” in In 2nd International Conference -2022(ICST2022), 2022, pp. 69–76, [Online]. Available: http://ir.lib.seu.ac.lk/handle/123456789/6340.

[18] S. Majumder and M. J. Deen, “Smartphone sensors for health monitoring and diagnosis,” Sensors (Switzerland), vol. 19, no. 9, pp. 1–45, 2019, doi: 10.3390/s19092164.

[19] A. Saboor, M. Usman, S. Ali, A. Samad, M. F. Abrar, and N. Ullah, “A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms,” Mobile Information Systems, vol. 2022, 2022, doi: 10.1155/2022/1410169.

[20] A. Gupta, R. Kumar, H. S. Arora, and B. Raman, “C-CADZ: computational intelligence system for coronary artery disease detection using Z-Alizadeh Sani dataset,” Applied Intelligence, vol. 52, no. 3, pp. 2436–2464, 2022, doi: 10.1007/s10489-021-02467-3.

[21] A. El Sheikh, N. Mahmoud, and A. Keshk, “Heart Disease Classification Based on Hybrid Ensemble Stacking Technique,” IJCI. International Journal of Computers and Information, vol. 8, no. 2, pp. 1–8, 2021, doi: 10.21608/ijci.2021.207732.